Enterprise Architecture
Executive Committee Update

March 19, 2015 | Thursday | 2:00-3:00 p.m. | Smith Center 1009
Agenda

• The Enterprise Architecture Vision
• EA Program Approach
• Business Needs and Benefits
• Current-State Analysis
  – Identity Data
  – SIS Data
  – Representative HR and Financial Data Flows
• Name and Address Information
  – Data Movement Protocols
  – Technologies Used
  – Directory Information Exchange Formats
• Interoperation
  – HBS Enterprise Data Exchange (HEDEX)
  – Domains
  – Vision and Proposed Guiding Principles
  – Establishing a Domain Work Plan
## The Enterprise Architecture Vision

### Our Vision for Enterprise Architecture

Provide a technology framework and a set of standards to enable acquisition, development, and deployment of IT services that maximize interoperation, minimize duplication, and simplify the IT environment across all of Harvard.

<table>
<thead>
<tr>
<th>Strategic Objectives</th>
<th>Guiding Principles</th>
<th>Key Performance Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Deliver an enterprise architecture framework that drives technology and development standards across Harvard</td>
<td>• Ensure that EA provides active direction and delivers value to the organization</td>
<td>• Decrease in project delivery timeframes to production</td>
</tr>
<tr>
<td>• Provide common approaches for integration across enterprise applications, processes, and data</td>
<td>• Counter complexity with common solutions</td>
<td>• Increase in the number of integrated applications using programmatic interfaces</td>
</tr>
<tr>
<td>• Align and rationalize technology decisions and investments</td>
<td>• Enable sharing of data across organizations</td>
<td>• Increase in the number of funded projects that conform to an EA checklist</td>
</tr>
<tr>
<td>• Identify redundant or conflicting processes and data across organizations</td>
<td>• Preference for open-source, COTS, and programmatic interfaces — both in what we obtain and what is produced</td>
<td>• Decrease in ad-hoc data sharing</td>
</tr>
<tr>
<td></td>
<td>• Encourage, define, and ultimately provide best-practice solutions</td>
<td>• Increase in automated data exchange</td>
</tr>
<tr>
<td></td>
<td>• Evolve framework and solutions with advances in technology</td>
<td>• Increase in the number of known authoritative data sources</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Decrease in the number of copies of data</td>
</tr>
</tbody>
</table>
EA Program Approach

Re-Evaluate: Identify Places Where EA Can Make an Impact

- Requirements and Needs
- Enterprise Technology Assessment
- Technology Trends and Best Practices

Enterprise Architecture Implementation Plan

Layers

Security
- User Experience
- Applications and Software Components
- Interoperation
- Data
- Middleware
- Infrastructure and IaaS
- Networking

Advisories, Methodologies, and Principles
- Architects
- UX Consultation
- Ad-Hoc Consultation
- ITCRB and PRC Reviews

Outreach and Training
- Evaluate Skills & Organizational Needs

Communication & Education
# Business Needs & Benefits

<table>
<thead>
<tr>
<th>Stakeholder</th>
<th>Experience Today</th>
<th>Imagine …</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Service Users: Faculty, Students, Staff</strong></td>
<td>• Time-consuming, repetitive entry of biographical and other data for each application/service   • Inability to answer simple business questions (“how many of X does Harvard have?”)   • Difficult and time-consuming to access data not already provided via an existing flow to a user’s organization</td>
<td>• Users do not have to re-enter data — apps get necessary info from a real-time service   • Faculty, students, and staff can ask questions on an ad-hoc basis and get reliable answers   • Faculty, students, and staff have flexible, timely access to data across the University</td>
</tr>
<tr>
<td><strong>IT Systems and Policy Decision-Makers</strong></td>
<td>• No guidance about interoperation requirements; all decisions about technology acquisition or development are local, which increases OP/CAPX   • No facility for cross-University data (“what do we mean by ‘address’ or ‘student’?”)</td>
<td>• Decision-makers have a simple interoperation requirements checklist to help them make local decisions that stick to Harvard-wide objectives   • We have a common understanding of what our data means across the University</td>
</tr>
<tr>
<td><strong>Systems and Software Engineering</strong></td>
<td>• Each integration is custom-coded, increasing technical risk   • So much time is spent on basic integration that tasks with greater user-perceived benefit are delayed</td>
<td>• Schools and departments have a set of reusable integration tools   • More efficient access to information frees up time to develop more user-valued services</td>
</tr>
<tr>
<td><strong>Operations</strong></td>
<td>• Multiple ad-hoc streams are time-consuming and brittle, creating create complexity for each system   • Difficult to identify faults and restore service due to multiple file transfer servers/flows with hard-to-understand interdependencies</td>
<td>• A centralized interoperation service available across Harvard, operated by a focused team   • No longer necessary to deploy, configure, and operate point-to-point integration services</td>
</tr>
</tbody>
</table>

Cost-efficient, flexible access to consistent information across the University that is easy for users and less complex for engineering and operational staff.
Current State: Identity Data

**University Financial Aid Liaison Office**
- SIS Data

**Campus Solutions**
- SIS +

**FAS Phone**
- DCE

**Multiple HBS Sources**
- HKS
- HLS
- Multiple HBS Locations
- GMAS
  - (and others - e.g, Library Campus Services, HUHS)

**Data Format**
- XML IAM Format
- XML HBS Format
- XML PeopleSoft New Hire Schema
- PeopleSoft Input XML
- DB View
- Human Input
- Pipe Delimited
- Oracle Datapump File
- Fixed Length Field Text File
- FindPerson/XML

**Transfer Protocol/Creation Method**
- Online forms/PeopleSoft Proprietary
- SFTP
- HTTPS Post
- RESTful FindPerson (JSON) service
- SQL query and SQLLOAD
- Web service
- RESTful FindPerson (JSON) service

**Data**
- Address, Person, Name, Email Address, Degree Data, Bio Data
- Address, Directory Listing Data, Job, Role, Employee Job Code, Department Code
- Name, DOB, last 4 of SSN, UUID, HUID, Gender
- Name, Person, Name Prefix, Middle Name, Last Name, Name Suffix
- 15 years service (T/F) (from secondary DB)
- Email Address, Directory Listing Data, Name, Person Address
- Employee File
- PeopleSoft
  - Person Type
  - Address Data
  - Degree Data
  - Bio Data
  - Name, DOB, last 4 of SSN, UUID, HUID, Gender
  - Name, Person, Name Prefix, Middle Name, Last Name, Name Suffix
  - 15 years service (T/F) (from secondary DB)

**Oracle Financial & Procurement**
- Employee File

**School Financial Aid Offices**
- School Data
Current State: SIS Data

SIS/Campus Solutions

Harvard School of Dental Medicine
- HLS (Graduate Programs)
- HLS (JD)
- HSPH
- Harvard Divinity School
- Harvard College
- GSAS

HBS (MBA)
HBS (Doctoral)
HMS

New Admit Bio Data

Name, DOB, last 4 of SSN, UUID, HUID, Gender

Name, DOB, last 4 of SSN

Name Address Data

Bio Data, Degree Data, Academic Progress

Find Person Service
CAADS (Monthly Feed)

New Admit Bio Data

HUID Assignment

Payroll Data

Bio Data, Workforce, Dept., Job, Location

PeopleSoft Payroll

PeopleSoft HR

GSAS
HKS
GSE
GSD

ESCI

Data Format
- JSON - (FindPerson)
- XML SIS Standard Admission XSD
- Tab Delimited Admission Data
- Fixed Length Field Text File
- Pipe Delimited
- PeopleSoft HR to SIS XML
- PeopleSoft Payroll to SIS XML

Transfer Protocol/Creation Method
- SOAP/XML
- SFTP into SIS
- SFTP from SIS
- Web REST/JSON - FindPerson service
Current State: Representative HR & Financial Data Flows

Data Format:
- Pipe Delimited Text
- Human Input
- Fixed Length Field Text File
- CSV Text File
- XML TOOR Feed
- XML Feed to GL
- Tab-Delimited Gift Data
- Direct Oracle Data
- Direct Oracle Data via Informatica

Transfer Protocol/Creation Method:
- SFTP
- Online forms/PeopleSoft Proprietary
- HTTPS Post
- Direct DB Feed
Name & Address: Data Movement Protocols

- Human Input Forms: 19%
- SQL query and SQL Load: 6%
- Web Service: 6%
- HTTP Post: 6%
- Restful Find Person: 6%
- SFTP: 50%
Name & Address: Technologies Used

- Unix Shell Script: 18%
- Perl: 6%
- Java: 6%
- Oracle PL/SQL: 6%
- Informatica: 12%
- Oracle SQL/LOAD & QUERY: 18%
- Oracle Datapump Utility: 18%
- On-Line Forms: 6%
- Oracle & Other Proprietary: 6%
Name & Address: Directory Information Exchange Formats

- XML IAM Format: 21%
- XML HBS Format: 16%
- XML PeopleSoft New Hire Schema: 11%
- PeopleSoft Input XML: 11%
- Oracle DB View: 11%
- Fixed Length Field Text File: 11%
- File Pipe Delimited: 5%
- Human Input: 5%
- Oracle Datapump File: 5%
- Find Person/JSON: 5%

XML IAM Format
XML PeopleSoft New Hire Schema
Oracle DB View
File Pipe Delimited
Fixed Length Field Text File
Human Input
Oracle Datapump File
Find Person/JSON
Case Study: HBS Enterprise Data Exchange (HEDEEx)

- Started in 2008; still evolving
- Uses standardized business objects to define main data sets under a publish-and-subscribe model
- Published data from source applications is transformed into standard formats
- Subscribed data is delivered in standard formats, with some exceptions
- Both changed data and full data sets are provided, some in near-real-time
- Design kept simpler by focusing only on HBS needs, resulting in a single-tenant solution
# Interoperation: Vision & Proposed Guiding Principles

## Our Vision for Interoperation in the Enterprise Architecture

Provide a framework and a set of standards to enable the acquisition, development, and deployment of integration services that maximize information sharing, minimize duplication, and simplify the IT environment across all of Harvard.

## Proposed Guiding Principles

- Similar data exchanged between applications must have standard definitions and formats
- Interoperation solutions will be implemented once
- Interoperation will use standardized mechanisms and protocols
- All exchanges of information must be authorized and protected appropriately
- Interoperation will support all solution environments
- Defining interoperation solutions will be open and transparent
Interoperation: Establishing a Domain Work Plan

- Architecture
  - Interoperation principles
  - Current-state survey and detailed assessments
  - Assist CIO Council initiatives: administration, academic, research
  - Identify ITCRB projects for embedded implementation (i.e. data warehouse Informatica strategy)
  - Advisories on data, transformation, and transport standards and practices
  - Reference architecture
  - Patterns (Transactional apps, data warehouse, SaaS & partner solutions, etc.)
  - Evaluate toolsets for best practice
  - Outreach and training

- Governance
  - Identify areas of need
  - Partner with leadership to identify governance approaches

Time
Thank you!